



XXI ENANCIB

Encontro Nacional de Pesquisa em Ciência da Informação

50 anos de Ciência da Informação no Brasil:
diversidade, saberes e transformação social

Rio de Janeiro • 25 a 29 de outubro de 2021

XXI Encontro Nacional de Pesquisa em Ciência da Informação – XXI ENANCIB

GT-8 – Informação e Tecnologia

METADADOS GEOESPACIAIS NA WEB SEMÂNTICA: UMA REVISÃO DA LITERATURA

GEOSPATIAL METADATA ON THE SEMANTIC WEB: A LITERATURE REVIEW

Clayton Martins Pereira, José Eduardo Santarem Segundo
{clayton.martins@unesp.br, santarem@usp.br}

Programa de Pós-Graduação em Ciência da Informação
Universidade Estadual Paulista (UNESP), Faculdade de Filosofia e Ciências, Marília

Modalidade: Trabalho Completo

Resumo: Metadados geoespaciais podem ser considerados uma especialização de metadados utilizados para a descrição de recursos de informação contendo uma localização espacial (dados geoespaciais). Por outro lado, a evolução da Web Semântica e o crescente uso de suas tecnologias e ferramentas, como o *Resource Description Framework* (RDF), e das melhores práticas de Dados Ligados (*Linked Data*) abriu caminho para se promover a publicação e interoperabilidade semântica dos metadados geoespaciais, os quais passaram a ser gerados e consumidos não somente por humanos, mas também por máquinas na Web. Diante disto, este trabalho tem como objetivo fazer um levantamento e seleção de artigos, indexados em bases de dados, sobre o tema 'publicação de metadados geoespaciais na Web Semântica'. Quanto ao enquadramento metodológico deste trabalho, trata-se de uma pesquisa de natureza aplicada, de abordagem qualitativa e quantitativa e do tipo descritiva, que emprega como métodos a revisão de literatura e a análise bibliométrica. Os procedimentos executados para a realização da revisão de literatura foram agrupados, de forma sequencial, ao longo de cinco etapas, as quais foram adaptadas das etapas 1, 2 e 3 do processo *ProKnow-C*, sendo ainda utilizada a equação *InOrdinatio* como ferramenta bibliométrica para ranqueamento das referências obtidas. Os resultados deste trabalho mostram que este tema ainda é pouco explorado na literatura, uma vez que, após a aplicação dos filtros de seleção e da análise dos textos, somente oito artigos atendiam aos critérios de análise estabelecidos. Além disso, foi possível identificar assuntos ainda não explorados na literatura sobre o tema de pesquisa.

Palavras-Chave: Metodologia. Revisão de literatura. Metadados geoespaciais. Web semântica.

Abstract: Geospatial metadata can be considered a specialization of metadata used for the description of information resources containing a spatial location (geospatial data). On the other hand, the evolution of the Semantic Web and the increasing use of its technologies and tools, such as the Resource Description Framework (RDF), and Linked Data best practices opened the way to promote the publication and semantic interoperability of geospatial metadata, which are now generated and consumed not only by humans, but also by machines on the Web. Therefore, this work aims to make a search and selection of articles, indexed in databases, on the theme 'publication of geospatial metadata on the Semantic Web'. The methodology of this work is an applied research of qualitative

and quantitative approach and of descriptive type, which uses as methods the literature review and bibliometric analysis. The procedures carried out for the literature review were grouped sequentially into five steps, which were adapted from steps 1, 2 and 3 of the ProKnow-C process, and also used the InOrdinatio equation as a bibliometric tool for ranking the references obtained. The results of this work show that this theme is still little explored in the literature, since, after applying the selection filters and analyzing the texts, only eight articles met the established analysis criteria. In addition, it was possible to identify subjects not yet explored in the literature about the research theme.

Keywords: Methodology. Literature Review. Geospatial metadata. Semantic web.

1 INTRODUÇÃO

A Web como conhecemos hoje ainda é uma Web de documentos, totalmente voltada para a leitura, extração e relacionamento de conteúdos por humanos, a qual foi formada a partir da criação crescente de *links* entre documentos de hipertexto, escritos em linguagem *HTML*, e da disponibilização de arquivos nos mais diversos formatos.

No entanto, ao longo das duas últimas décadas, os usuários começaram a fazer upload de conteúdo textual e de multimídia a taxas cada vez mais altas, o que somado ao enorme volume de páginas Web e o maior número de requisições, passou a exigir dos aplicativos Web novas maneiras de lidar com documentos, de forma que máquinas pudessem entender com quais dados estavam lidando. Isso levou ao surgimento da Web 3.0, chamada de Web Semântica ou Web de Dados, na qual a ideia principal é fornecer um contexto aos documentos na Web em um formato legível por máquina (PO *et al.*, 2020).

A partir da Web Semântica, passou-se a estimular a publicação de conteúdo junto com os metadados que os descrevem de uma forma compreensível por máquina. Estas descrições legíveis por máquina permitem aos gerenciadores de conteúdo adicionarem significado ao conteúdo, de forma que uma máquina pode processar o próprio conhecimento, em vez de texto, obtendo resultados mais significativos e ajudando os computadores a realizar a coleta e pesquisa automatizada de informações (PO *et al.*, 2020).

Neste contexto, os metadados geoespaciais podem ser considerados uma especialização de metadados, utilizados para a descrição de recursos de informação contendo dados geoespaciais. Os dados geoespaciais (também chamados de dados geográficos) caracterizam-se pela localização espacial (ou geográfica) onde um determinado fenômeno ocorre, acompanhada de atributos que descrevem esse fenômeno. Esta localização espacial deve ter como sistema de referência a superfície da Terra, sendo geralmente empregados sistemas de coordenadas em sua descrição (W3C; OGC, 2017).

A evolução da Web Semântica e o crescente uso de suas tecnologias e ferramentas, como o *Resource Description Framework* (RDF), e das melhores práticas de Dados Ligados (*Linked Data*) propiciou, de acordo com Pereira, Ferneda e Santarem Segundo (2020, p. 95), “...o surgimento de iniciativas de abertura de dados por diversos governos e instituições”, o que deu origem ao chamado paradigma dos Dados Abertos Ligados (*Linked Open Data*). Tal paradigma abriu caminho para se promover a publicação e interoperabilidade semântica dos metadados geoespaciais, os quais passaram a ser gerados e consumidos não somente por humanos, mas também por máquinas e agentes de software na Web.

Diante disto, este trabalho tem como objetivo fazer um levantamento e seleção de artigos publicados em periódicos, ou anais de conferências, indexados em bases de dados, sobre o tema: publicação de metadados geoespaciais na Web Semântica. Quanto ao enquadramento metodológico deste trabalho, trata-se de uma pesquisa de natureza aplicada, de abordagem qualitativa e quantitativa e do tipo descritiva, que emprega como métodos a revisão de literatura e a análise bibliométrica.

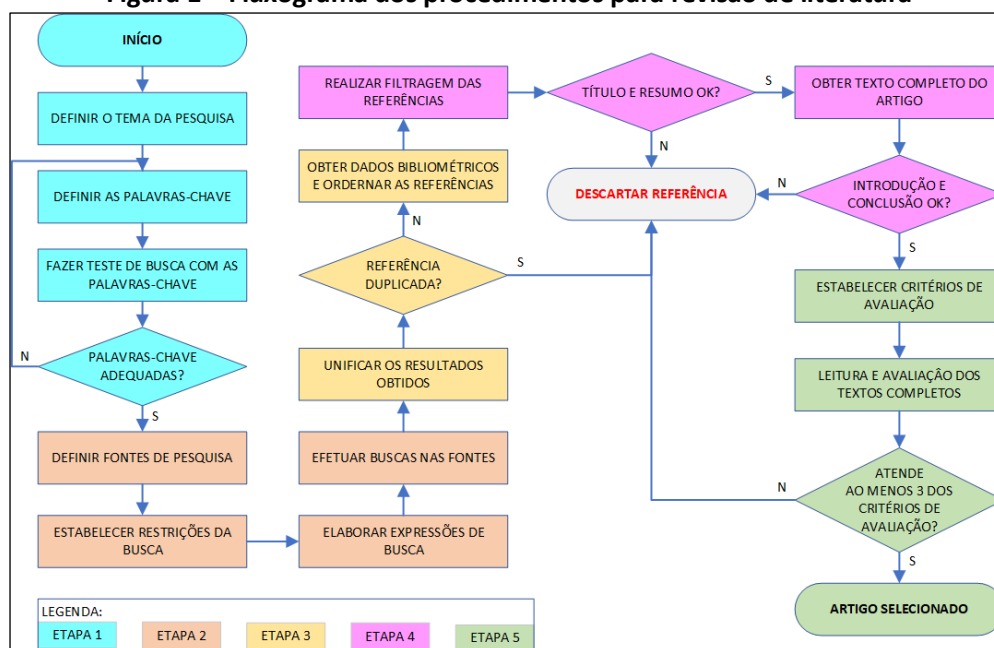
Nas seções a seguir, são descritos os procedimentos metodológicos desta revisão de literatura, bem como apresentados e analisados os resultados obtidos, seguidos das considerações finais.

2 PROCEDIMENTOS METODOLÓGICOS

A metodologia de execução desta revisão de literatura tem como base o processo *ProKnow-C* que, de acordo com Afonso *et al.* (2012, p. 49), “se constitui em uma metodologia para construção do conhecimento”, a qual é composta por quatro etapas: “1) seleção do portfólio bibliográfico que proporcionará a revisão de literatura; 2) análise bibliométrica do portfólio bibliográfico; 3) análise sistêmica do portfólio bibliográfico; 4) elaboração dos objetivos de pesquisa.”. Estas etapas consistem em uma série de procedimentos sequenciais pré-estabelecidos, que vão desde a definição das fontes de busca até a seleção do portfólio relevante sobre o tema de pesquisa (AFONSO *et al.*, 2012).

Desta forma, todos os procedimentos executados para a realização da revisão de literatura apresentada neste trabalho foram agrupados, de forma sequencial, ao longo de cinco etapas, adaptadas das etapas 1, 2 e 3 do processo *ProKnow-C* (AFONSO *et al.*, 2012). A descrição dos procedimentos e a ordem de execução dos mesmos, bem como o sequenciamento das etapas (diferenciadas por cores) estão representados no fluxograma apresentado na Figura 1.

Figura 1 – Fluxograma dos procedimentos para revisão de literatura



Fonte: Adaptado de Afonso *et al.* (2012).

Nas subseções a seguir, são detalhadas cada uma das etapas, e respectivos procedimentos, apresentadas no fluxograma da Figura 1.

2.1 Etapa 1: Definição do tema da pesquisa e das palavras-chave relacionadas

Na primeira etapa da revisão de literatura (cor ciano no fluxograma), foram realizados inicialmente os procedimentos de definição do tema de pesquisa e, a partir deste tema, de escolha das palavras-chave que o representem adequadamente. No entanto, este procedimento de escolha de palavras-chave requer plena atenção do pesquisador, uma vez que determinará quais artigos serão selecionados para as etapas posteriores da revisão de literatura. Caso as palavras-chave não sejam escolhidas adequadamente, alguns artigos relevantes sobre o tema de pesquisa podem ficar de fora quando da realização do procedimento de busca nas bases de dados, o que irá impactar negativamente a qualidade da revisão de literatura realizada e, consequentemente, na definição do estado da arte do tema da pesquisa.

Assim, para o tema de pesquisa ‘Publicação de metadados geoespaciais na Web Semântica’, foram definidas as seguintes palavras-chave: Metadados geoespaciais (*Geospatial metadata*); Web Semântica (*Semantic Web*); Web de dados (*Web of data*); Dados abertos ligados (*Linked Open Data*); Framework para descrição de recursos (*Resource Description Framework*); e, Informação Geoespacial (*Geospatial information*).

Escolhidas as palavras-chave iniciais, no procedimento seguinte, as traduções para o inglês das mesmas foram testadas, tanto isoladamente quanto de maneira combinada, nos

seguintes mecanismos de busca: *Scopus* e *Web of Science*. A partir dos resultados obtidos nestas buscas, a palavra-chave '*geospatial metadata*' foi substituída por '*geo* metadata*', onde o caractere '*' representa 'qualquer variação', ou seja, abrangerá tanto os termos '*geospatial*' como '*geographic*', pois ambos costumam ser usados indistintamente para se referir a este tipo de metadado. Da mesma forma, a palavra-chave '*linked data*' foi incluída, uma vez que alguns artigos utilizam este termo por considerar que nem todo dado ligado (*linked data*) é aberto (*open*). Por sua vez, a palavra-chave '*Resource Description Framework*' foi substituída por seu acrônimo '*RDF*', mais comumente utilizado nos artigos, e a palavra-chave '*geospatial information*', por não trazer resultados relevantes, foi excluída.

2.2 Etapa 2: Busca em bases de dados indexadas

Após a definição do tema de pesquisa e a escolha e teste das palavras-chave que o representam, na segunda etapa (cor bege no fluxograma) foram definidas as fontes de busca de artigos, e as restrições da busca nessas fontes, tais como os tipos de documentos a serem encontrados, o período de abrangência e as respectivas expressões de busca, elaboradas a partir das palavras-chave definidas na etapa anterior. O Quadro 1 apresenta as fontes e as restrições de busca estipulados para este trabalho.

Quadro 1 – Definição das fontes e das restrições das buscas a serem efetuadas.

FONTES DE BUSCA	– <i>Scopus</i> – <i>Web of Science</i>
TIPOS DE DOCUMENTOS	– Artigos em periódicos – Artigos em conferência
PERÍODO DE ABRANGÊNCIA	– Últimos 10 anos (2011 a 2020/21)
EXPRESSÕES DE BUSCA	– Scopus: TITLE-ABS-KEY("geo* metadata") AND TITLE-ABS-KEY("semantic web" OR "web of data" OR "linked data" OR "linked open data" OR "rdf") AND DOCTYPE(ar OR cp) AND PUBYEAR > 2010 – Web of Science: TS=("geo* metadata" AND ("semantic web" OR "web of data" OR "linked data" OR "linked open data" OR "rdf")) AND DT=(Article OR Proceedings Paper) AND PY=(2011-2021)

Fonte: Elaborado pelos autores.

Definidas as fontes e as restrições da busca, os resultados obtidos em cada uma das bases consultadas foram exportados para arquivos no formato *BibTeX* que foram salvos para serem organizados na etapa seguinte.

2.3 Etapa 3: Organização dos resultados da busca

Na terceira etapa (cor amarela no fluxograma), os resultados obtidos nas buscas efetuadas na etapa anterior foram organizados com o intuito de facilitar a execução das etapas seguintes, de filtragem e de seleção de textos. Para isso, foram realizados os seguintes passos:

1) Unificar os metadados coletados na *Scopus* e *Web of Science*, descartando as referências duplicadas (script em linguagem R); e, 2) Obter dados bibliométricos e ordenar as referências aplicando a equação *InOrdinatio* (Ano+Citações+Fator de Impacto).

No primeiro passo desta etapa, os dois arquivos (em formato *BibTeX*) contendo os resultados das buscas efetuadas nas bases *Scopus* e *Web of Science* foram unificados, sendo removidas as referências em duplicidade. Os dados unificados foram então salvos em um arquivo csv (texto separado por vírgulas, em português), o qual pode ser facilmente convertido em planilha para ser usada no passo seguinte. Para a execução deste passo foi elaborado um script em linguagem R, por meio do software *RStudio*, o qual é apresentado no Quadro 2.

Quadro 2 – Script, em linguagem R, para unificação e remoção de resultados duplicados das buscas

```
library(bibliometrix)
x <- convert2df("path_to_file/scopus.bib", dbsource = "scopus", format = "bibtex")
y <- convert2df("path_to_file/wos.bib", dbsource = "isi", format = "bibtex")
z <- mergeDbSources(X, Y, remove.duplicated = TRUE)
w <- z[,c("AU", "TI", "SO", "AB", "DE", "ID", "DI", "LA", "DT", "TC", "PY")]
write.table(w, "path_to_file/artigos.csv", sep=";", row.names=FALSE)
```

Fonte: Elaborado pelos autores.

No script apresentado, é utilizada uma biblioteca (*library*) chamada *bibliometrix* (ARIA; CUCCURULLO, 2017). Por meio desta biblioteca é que são lidos os arquivos de resultados, em formato *BibTeX*, obtidos nas buscas nas bases *Scopus* (variável X) e *Web of Science* (variável Y), e efetuada a unificação dos mesmos, removendo-se as duplicidades (variável Z). Por fim, são selecionadas e organizadas as colunas autor (AU), título (TI), nome do periódico ou dos anais de evento (SO), resumo (AB), palavras-chave do autor (DE), palavras-chave da indexação (ID), identificador DOI (DI), idioma (LA), tipo de artigo (DT), total de citações (TC) e ano de publicação (PY), as quais são gravadas em um arquivo no formato csv (*write.table*).

No segundo passo foram coletados dados bibliométricos das referências unificadas no primeiro passo, mais especificamente os valores de Fator de Impacto (FI ou *IF*, em inglês) de cada periódico. As fontes de consulta mais comumente utilizadas para obter o FI são o *Scimago Journal Rank (SJR)*, publicado pela *Scopus*, o *Journal Citation Report (JCR)*, publicado pela *Clarivate (Web of Science)*, e o *Google Scholar H Index*. Os valores de FI obtidos para cada referência foram inseridos em uma nova coluna acrescentada às existentes no arquivo gerado.

A obtenção dos valores de FI dos periódicos se fez necessária neste passo para possibilitar a ordenação dos resultados, por meio da aplicação da equação *InOrdinatio* (PAGANI; KOVALESKI; RESENDE, 2015). A equação *InOrdinatio* é produto de um modelo multicritério de

avaliação, utilizado para ranquear artigos científicos, baseado na combinação de três critérios: a) o fator de impacto do periódico em que o artigo foi publicado; b) o número de citações; e c) o ano de publicação. O Quadro 3 apresenta a equação e a descrição de suas variáveis.

Quadro 3 – Equação *InOrdinatio* para ranquear artigos científicos

$InOrdinatio = (IF/1000) + \alpha * [10 - (RY - PY)] + (\Sigma Ci)$, onde:

IF = *Impact Factor*: fator de impacto do periódico no qual o artigo foi publicado;

α = Peso a ser dado ao critério “ano de publicação”, podendo variar de 1 a 10;

RY = *Research Year*: ano em que a pesquisa bibliográfica está sendo feita;

PY = *Publication Year*: ano de publicação do artigo analisado;

ΣCi = *Citation*: número de citações do artigo analisado.

Fonte: Adaptado de Pagani, Kovaleski e Resende (2015, p. 2120).

Na equação apresentada, o fator do impacto (IF) do periódico é dividido por 1.000, com o intuito de normalizar seu valor com relação ao demais critérios (ano de publicação e número de citações). Por sua vez, o critério ano é calculado pela diferença entre o ano em que a pesquisa bibliográfica está sendo realizada (RY) e o ano de publicação do artigo (PY), sendo o resultado dessa diferença (RY-PY) subtraído da constante 10, que representa o horizonte máximo da pesquisa bibliográfica realizada (últimos 10 anos), e o resultado obtido é multiplicado pelo peso (α , de 1 a 10) dado a este critério em relação aos demais. Por fim, o critério número de citações (ΣCi) é a reprodução do valor apresentado na coluna TC, derivada dos resultados das buscas efetuadas. O valor final da equação é a soma dos valores resultantes desses três critérios.

Assim, obtido o valor do *InOrdinatio* de cada referência, foi realizada a ordenação (ranqueamento) das referências, partindo do maior para o menor valor desta coluna, o que auxilia na realização da etapa 4 desta revisão de literatura, que é a aplicação de filtros para seleção dos textos a serem analisados. Da mesma forma, nos casos de buscas que retornem um número muito grande de resultados, o valor do *InOrdinatio* pode servir de ‘nota de corte’, de forma a garantir que os artigos de maior relevância para a revisão de literatura sejam analisados.

2.4 Etapa 4: Aplicação de filtros para seleção de textos

Na quarta etapa desta revisão de literatura (cor magenta no fluxograma) foi realizada a filtragem das referências obtidas nas buscas efetuadas nas bases de dados, após os procedimentos de unificação e ordenação realizados na etapa anterior. Como resultado desta etapa, foram obtidos os textos completos dos artigos a serem analisados na etapa seguinte desta revisão de literatura. A aplicação de filtros foi realizada a partir dos três passos: 1) Análise de Título e Resumo; 2) Obter texto completo do artigo; e, 3) Análise de Introdução e Conclusão.

O primeiro filtro aplicado nesta etapa foi a análise do título e do resumo de cada um dos artigos. Caso estes dois elementos do artigo estivessem alinhados ao tema de pesquisa, então seguia-se para o segundo passo, que é o de obtenção do texto completo do artigo, caso contrário, a referência era imediatamente descartada. Para auxiliar a execução do passo 2, de obtenção dos textos completos, foi utilizado o gerenciador de referências *Zotero*, o que possibilitou a organização dos textos baixados, e respectivas referências completas, bem como facilitar o processo de elaboração de referências bibliográficas.

Após a obtenção dos textos completos dos artigos selecionados, foi realizada, no terceiro passo, a análise das seções de introdução e de conclusão de cada artigo. Da mesma forma que no filtro anterior, caso estas duas seções do artigo estivessem alinhadas ao tema de pesquisa, então o artigo era selecionado para a próxima etapa da revisão de literatura e, caso contrário, a referência e o respectivo texto completo eram descartados.

2.5 Etapa 5: Análise dos textos completos

Na última etapa (cor verde no fluxograma) foi realizada a análise do texto completo de cada um dos artigos selecionados na etapa anterior. Para isso, foram executados os seguintes passos: 1) Estabelecer os critérios de avaliação dos artigos; e, 2) Leitura e avaliação dos textos completos dos artigos. Como passo inicial desta etapa, foram estabelecidos os critérios de avaliação dos artigos selecionados para terem os respectivos textos completos analisados. Estes critérios estão descritos no Quadro 4.

Quadro 4 – Critérios para avaliação dos artigos

ID	CRITÉRIO
1	Propõe processo, método ou técnica para a publicação de metadados (geo)espaciais na Web?
2	Aborda conceitos e/ou emprega ferramentas da Web Semântica aplicados a metadados (geo)espaciais?
3	Detalha pelo menos um dos seguintes procedimentos para publicação de metadados espaciais na Web? a) Conversão de formatos (ISO 19115/GML/XML) para RDF; b) Armazenamento de triplas RDF (<i>Triple Store</i>); c) Disponibilização e Acesso na Web de Dados;
4	Aborda a aplicação de boas práticas para publicação de dados na Web (<i>Linked Data</i> e/ou <i>W3C/OGC</i>);
5	Apresenta metodologia bem definida e resultados condizentes com os objetivos propostos?

Fonte: Elaborado pelos autores.

Uma vez estabelecidos os critérios para avaliação dos artigos, no passo seguinte os textos completos dos artigos selecionados para esta etapa foram lidos integralmente e avaliados quanto aos critérios estabelecidos. Ao final, os artigos que atenderam pelo menos três dos cinco critérios foram selecionados para compor o portfólio da pesquisa.

Na seção a seguir são apresentados e analisados os resultados obtidos da aplicação de cada uma das etapas e passos descritos nesta seção.

3 ANÁLISE DOS RESULTADOS

A partir do tema de pesquisa e das respectivas palavras-chave que o representam, definidos na etapa 1, e das fontes de pesquisa, das restrições e das expressões de busca (vide Quadro 1), definidas na etapa 2, foram realizadas as buscas nas bases de dados *Scopus* e *Web of Science*, sendo encontradas 21 referências na primeira base e 11 referências na segunda base.

Durante o procedimento de unificação dos resultados (passo 1 da etapa 3), utilizando o *script* em linguagem *R*, foi possível notar que 9 das 11 referências encontradas na base de dados *Web of Science* também foram encontradas na base de dados *Scopus*, e que uma das referências encontradas nesta última estava duplicada. Assim, foram obtidas 22 referências únicas, as quais foram salvas em um único arquivo no formato *csv*.

Já no passo 2 da etapa 3, foram coletados, por meio de consulta ao *Scimago Journal Rank (SJR)*¹, quando encontrados, os valores do Fator de Impacto dos periódicos ou dos anais de eventos, conforme o tipo de artigo de cada referência. O *SJR* foi escolhido como fonte de consulta para o FI dos artigos selecionados neste trabalho devido ao fato de que 20 das 22 referências encontradas nas buscas estavam indexadas na base de dados *Scopus*.

Assim, coletado o valor do FI de cada referência, foi calculado o respectivo valor da equação *InOrdinatio*, o que possibilitou a ordenação (ranqueamento) das referências obtidas na etapa anterior. Na equação, foi atribuído o valor 1 para a variável α , o que significa que o ano de publicação do artigo tem o mesmo peso que os demais critérios (fator de impacto e total de citações). Para a anotação destes dois indicadores bibliométricos foram adicionadas duas novas colunas (Fator de Impacto e *InOrdinatio*) na planilha.

Na etapa 4 foram aplicados os filtros de seleção de textos, sendo o primeiro deles a leitura do título e resumo de cada referência. A Tabela 1 mostra os dados das colunas 'Autores', 'Ano de publicação', 'Tipo de documento', 'Total de citações', 'Fator de Impacto' e '*InOrdinatio*', ordenada por esta última coluna, além de mostrar as referências selecionadas, após a leitura do título e resumo, para as quais devem ser obtidos os respectivos textos completos.

Tabela 1 – Relação das referências obtidas e aplicação dos filtros de seleção de textos

Autores (AU)	Ano de Publicação (PY)	Tipo de Documento (DT)	Total de Citações (TC)	Fator de Impacto (IF/1000)	<i>InOrdinatio</i> ($\alpha = 1$)	Título ok?	Resumo ok?	Obter texto completo?
YUAN <i>et al.</i>	2013	Article	15	2,616	19,62	Sim	Sim	Sim
KALANTARI <i>et al.</i>	2014	Article	12	1,356	16,36	Sim	Não	Não
LI <i>et al.</i>	2020	Conference Paper	0	4,027	13,03	Sim	Sim	Sim
CHEATHAM <i>et al.</i>	2018	Article	6	-	13,00	Sim	Sim	Sim

¹ Disponível em <https://www.scimagojr.com/journalrank.php>. Acesso via Portal de Periódicos CAPES.

Autores (AU)	Ano de Publicação (PY)	Tipo de Documento (DT)	Total de Citações (TC)	Fator de Impacto (IF/1000)	InOrdinatio ($\alpha = 1$)	Título ok?	Resumo ok?	Obter texto completo?
YUE <i>et al.</i>	2012	Article	11	0,491	12,49	Sim	Sim	Sim
FUGAZZA <i>et al.</i>	2018	Article	3	0,655	10,66	Sim	Sim	Sim
FUGAZZA <i>et al.</i>	2016a	Article	5	0,655	10,66	Sim	Sim	Sim
FUGAZZA <i>et al.</i>	2021	Article	0	0,386	10,39	Sim	Sim	Sim
GRAUPNER e NÜST	2020	Conference Paper	0	0,537	9,54	Não	Não	Não
FUGAZZA <i>et al.</i>	2016b	Conference Paper	4	0,175	9,18	Sim	Sim	Sim
MCKEE, G.	2019	Article	0	0,771	8,77	Não	Não	Não
FUGAZZA <i>et al.</i>	2017	Article	2	0,655	8,66	Sim	Sim	Sim
SHVAIKO <i>et al.</i>	2012	Conference Paper	7	-	8,00	Sim	Sim	Sim
ZHANG <i>et al.</i>	2013	Conference Paper	2	-	4,00	Sim	Sim	Sim
MENEZES <i>et al.</i>	2014	Article	0	0,772	3,77	Não	Não	Não
RENTERIA-AGUALIMPIA <i>et al.</i>	2013	Conference Paper	1	0,200	3,20	Sim	Sim	Sim
DO <i>et al.</i>	2014	Conference Paper	0	0,177	3,18	Não	Não	Não
FERNANDEZ-ORDONEZ e CAROLINA	2014	Conference Paper	0	-	3,00	Sim	Sim	Sim
GIANNOPOULOS <i>et al.</i>	2014	Conference Paper	0	-	3,00	Não	Não	Não
RITSCHER <i>et al.</i>	2012	Article	1	0,309	2,31	Sim	Não	Não
BAPTISTA <i>et al.</i>	2013	Conference Paper	0	0,117	2,12	Não	Não	Não
AKLI-ASTOUATI, GUEBAILI e MOKHTARI	2011	Article	0	0,111	0,11	Sim	Sim	Sim

Fonte: Elaborado pelos autores.

Nesta tabela, é possível perceber que, mesmo dando ao ano de publicação o mesmo peso dado aos critérios ‘fator de impacto’ e ‘total de citações’, a maioria dos artigos mais antigos e menos citados ficaram posicionados na metade inferior da mesma. Isso pode indicar que a equação tenha conseguido ordenar as referências de acordo com sua relevância científica para o tema de pesquisa deste trabalho, o que será validado ao término da etapa 5. Também é possível verificar na tabela que, na etapa 4, foram descartadas oito referências devido ao fato dos respectivos títulos e/ou resumos não estarem alinhados ao tema de pesquisa.

Após a obtenção dos textos completos dos artigos selecionados no passo 2, foi realizada, no passo 3, a leitura das seções de introdução e resumo de cada artigo. A Tabela 2 mostra, apenas para os artigos selecionados no passo 1, os resultados deste procedimento.

Tabela 2 – Aplicação dos filtros de seleção de textos do Passo 3 da Etapa 4

Autores (AU)	Ano de Publicação (PY)	Tipo de Documento (DT)	Total de Citações (TC)	Fator de Impacto (IF/1000)	InOrdinatio ($\alpha = 1$)	Introdução e Conclusão ok?
YUAN <i>et al.</i>	2013	Article	15	2,616	19,62	Sim
LI <i>et al.</i>	2020	Conference Paper	0	4,027	13,03	Sim
CHEATHAM <i>et al.</i>	2018	Article	6	-	13,00	Sim
YUE <i>et al.</i>	2012	Article	11	0,491	12,49	Sim
FUGAZZA <i>et al.</i>	2018	Article	3	0,655	10,66	Sim
FUGAZZA <i>et al.</i>	2016a	Article	5	0,655	10,66	Sim
FUGAZZA <i>et al.</i>	2021	Article	0	0,386	10,39	Sim
FUGAZZA <i>et al.</i>	2016b	Conference Paper	4	0,175	9,18	Sim
FUGAZZA <i>et al.</i>	2017	Article	2	0,655	8,66	Sim
SHVAIKO <i>et al.</i>	2012	Conference Paper	7	-	8,00	Sim
ZHANG <i>et al.</i>	2013	Conference Paper	2	-	4,00	Sim
RENTERIA-AGUALIMPIA <i>et al.</i>	2013	Conference Paper	1	0,200	3,20	Sim
FERNANDEZ-ORDONEZ e CAROLINA	2014	Conference Paper	0	-	3,00	Não
AKLI-ASTOUATI, GUEBAILI e MOKHTARI	2011	Article	0	0,111	0,11	Sim

Fonte: Elaborado pelos autores.

Ao observar a tabela, é possível notar que da aplicação do procedimento de leitura das seções de introdução e conclusão de cada artigo, foi descartada mais uma referência, por não estar alinhada ao tema de pesquisa.

Por fim, na Etapa 5, foi efetuada a leitura integral dos textos selecionados na etapa anterior. Para cada texto lido, foram respondidos, no Passo 2, os critérios de análise estabelecidos no Passo 1 desta etapa (vide Quadro 4). A Tabela 3 apresenta os resultados da análise de cada texto (Passo 2), indicando os que foram selecionados, por atender ao menos três dos critérios de avaliação, para compor o portfólio da pesquisa.

Tabela 3 – Resultados da leitura e análise dos textos realizada na Etapa 5

Autores (AU)	Ano de Publicação (PY)	Critério 1?	Critério 2?	Critério 3?				Critério 4?	Critério 5?	Artigo Selecionado?
				questão a)	questão b)	questão c)	atende critério			
YUAN <i>et al.</i>	2013	SIM	SIM	Sim	Não	Sim	SIM	NÃO	SIM	SIM
LI <i>et al.</i>	2020	NÃO	SIM	Não	Não	Não	NÃO	NÃO	SIM	NÃO
CHEATHAM <i>et al.</i>	2018	NÃO	SIM	Sim	Não	Sim	SIM	SIM	SIM	SIM
YUE <i>et al.</i>	2012	NÃO	SIM	Sim	Não	Não	SIM	NÃO	SIM	SIM
FUGAZZA <i>et al.</i>	2018	NÃO	SIM	Sim	Não	Sim	SIM	NÃO	SIM	SIM
FUGAZZA <i>et al.</i>	2016a	NÃO	SIM	Não	Não	Não	NÃO	NÃO	NÃO	NÃO
FUGAZZA <i>et al.</i>	2021	SIM	SIM	Sim	Não	Sim	SIM	NÃO	SIM	SIM
FUGAZZA <i>et al.</i>	2016b	NÃO	SIM	Sim	Não	Sim	SIM	NÃO	SIM	SIM
FUGAZZA <i>et al.</i>	2017	NÃO	SIM	Sim	Não	Sim	SIM	NÃO	SIM	SIM
SHVAIKO <i>et al.</i>	2012	SIM	SIM	Não	Não	Sim	SIM	NÃO	SIM	SIM
ZHANG <i>et al.</i>	2013	NÃO	SIM	Não	Não	Não	NÃO	NÃO	SIM	NÃO
RENTERIA-AGUALIMPIA <i>et al.</i>	2013	NÃO	NÃO	Não	Não	Não	NÃO	NÃO	SIM	NÃO
AKLI-ASTOUATI, GUEBAILI e MOKHTARI	2011	NÃO	SIM	Não	Não	Não	NÃO	NÃO	NÃO	NÃO

Fonte: Elaborado pelos autores.

A partir dos dados apresentados nesta tabela, é possível identificar que nesta etapa foram descartadas outras cinco referências, por não terem atendido a, pelo menos, três dos critérios estabelecidos para análise dos textos. Também é possível notar, quanto ao atendimento individual dos critérios e quesitos (critério 3) avaliados, que: a) somente três dos artigos analisados atenderam ao critério 1 - propor processo, método ou técnica para a publicação de metadados (geo)espaciais na Web, tema da pesquisa; b) nenhum artigo atendeu ao quesito b - armazenamento de triplas RDF, do critério 3 - detalhar pelo menos um dos procedimentos (quesitos) para publicação de metadados (geo)espaciais na Web; e, c) somente um artigo atendeu ao critério 4 - abordar a aplicação de boas práticas para publicação de dados na Web. Isto mostra que o tema de pesquisa ainda é pouco explorado na literatura, o que pode abrir um leque de questões relacionadas ao tema a serem estudadas nos próximos anos.

Assim, como resultado final desta revisão de literatura, é apresentada a Tabela 4, que relaciona os artigos escolhidos para compor o portfólio da pesquisa sobre o tema 'Publicação de metadados (geo)espaciais na Web Semântica'.

Tabela 4 – Artigos selecionados para compor o portfólio da pesquisa

Autores (AU)	Ano de Publicação (PY)	Tipo de Documento (DT)	Total de Citações (TC)	Fator de Impacto (IF/1000)	<i>InOrdinatio</i> ($\alpha = 1$)
YUAN <i>et al.</i>	2013	Article	15	2,616	19,62
CHEATHAM <i>et al.</i>	2018	Article	6	-	13,00
YUE <i>et al.</i>	2012	Article	11	0,491	12,49
FUGAZZA <i>et al.</i>	2018	Article	3	0,655	10,66
FUGAZZA <i>et al.</i>	2021	Article	0	0,386	10,39
FUGAZZA <i>et al.</i>	2016b	Conference Paper	4	0,175	9,18
FUGAZZA <i>et al.</i>	2017	Article	2	0,655	8,66
SHVAIKO <i>et al.</i>	2012	Conference Paper	7	-	8,00

Fonte: Elaborado pelos autores.

Com relação à utilização da equação *InOrdinatio* (PAGANI; KOVALESKI; RESENDE, 2015) para ordenar (ranquear) as referências obtidas nas buscas efetuadas, ao visualizar a Tabela 5 é possível verificar que todos os artigos apresentaram valor calculado de *InOrdinatio* $\geq 8,00$, ou seja, todos os artigos escolhidos estavam na metade superior da listagem ordenada. Assim, caso fosse adotado este valor de *InOrdinatio* como ‘nota de corte’ (filtro) para descarte de artigos logo no início da Etapa 4, o resultado deste processo de seleção de artigos atingiria o mesmo resultado, porém haveria uma quantidade menor de artigos a serem avaliados por meio da leitura do texto completo, sem risco de perda da qualidade do trabalho.

4 CONSIDERAÇÕES FINAIS

Neste trabalho foram descritos de forma detalhada todos os procedimentos metodológicos, e os respectivos resultados obtidos, de uma revisão de literatura para compor o portfólio de artigos a serem utilizados para estabelecer o estado-da-arte sobre o tema ‘publicação de metadados (geo)espaciais na Web Semântica’.

Os resultados deste trabalho mostraram que este tema é pouco explorado na literatura, uma vez que a consulta nas bases de dados *Scopus* e *Web of Science* retornaram uma quantidade pequena de referências únicas (somente 22 artigos) para as palavras-chave utilizadas nas expressões de busca. Destas, após a aplicação dos filtros de seleção e da análise dos textos, somente oito atendiam aos critérios de avaliação estabelecidos e estavam plenamente alinhados ao tema de pesquisa.

No entanto, os critérios de análise estabelecidos permitiram identificar assuntos ainda não encontrados na literatura sobre o tema de pesquisa, tais como as técnicas e ferramentas para armazenamento de triplas RDF de metadados geoespaciais, que não foi abordado por nenhum dos artigos, e a aplicação de boas práticas para os metadados geoespaciais na Web, que foi abordado brevemente em apenas um dos artigos analisados.

Além disso, este trabalho mostrou que a aplicação da equação *InOrdinatio* para a organização (ranqueamento) das referências encontradas nas buscas, ao levar em conta o fator de impacto dos periódicos e anais de conferências, o ano de publicação do artigo, e o total de citações, se mostrou eficaz ao classificar as referências encontradas por sua relevância científica, o que foi comprovado pelo fato de que ao final da última etapa desta revisão de literatura, os artigos selecionados foram justamente aqueles posicionados na metade superior da lista de referências, ou seja, aquelas que apresentaram maior valor de *InOrdinatio*.

Diante disto, a aplicação de tal equação tem potencial para otimizar o trabalho em processos de revisão de literatura ou seleção de referências com um elevado número de resultados, ao possibilitar ao pesquisador definir um 'valor de corte' do *InOrdinatio*, onde as referências com valor abaixo deste 'valor de corte' poderiam ser descartadas sem afetar de forma significativa a qualidade de todo o processo. Assim, o pesquisador pode dispendar seu tempo na análise de texto somente das referências com maior relevância científica, onde estarão concentrados os artigos que serão selecionados no processo.

Como trabalhos futuros, a revisão de literatura apresentada neste trabalho será agregada a um projeto de pesquisa em andamento sobre o tema, conduzido pelos autores. Com relação aos procedimentos metodológicos apresentados (fluxograma) recomenda-se aplicá-lo a uma revisão de literatura com um elevado número de referências encontradas nas buscas efetuadas, de forma a validar a utilização da equação *InOrdinatio* em substituição à etapa 4 apresentada neste trabalho.

REFERÊNCIAS

- AFONSO, M. H. F.; SOUZA, J. V. de; ENSSLIN, S. R.; ENSSLIN, L. Como construir conhecimento sobre o tema de pesquisa? Aplicação do processo PROKNOW-C na busca de literatura sobre avaliação do desenvolvimento sustentável. **Revista de Gestão Social e Ambiental**, São Paulo, v. 5, n. 2, p. 47-62, mai./ago. 2012. DOI: 10.5773/rgsa.v5i2.424.
- AKLI-ASTOUATI, K.; GUEBAILI, R.; MOKHTARI, A. Geographic metadata: data quality, uncertainty and imprecision. **International Journal of Reasoning-based Intelligent Systems**, Geneva, v. 3, n. 3–4, p. 164–172, 2011. DOI: 10.1504/IJRIS.2011.043541.
- ARIA, M.; CUCCURULLO, C. Bibliometrix: An R-tool for comprehensive science mapping analysis. **Journal of Informetrics**, Amsterdam, v. 11, n. 4, p. 959–975, 2017. DOI: 10.1016/j.joi.2017.08.007.
- BAPTISTA, C. S.; MENEZES, L. C.; OLIVEIRA, M. G.; FALCÃO, A. G. R.; MARINHO, L. B. On building semantically enhanced location-based social networks. In: SIMPÓSIO BRASILEIRO DE GEO-INFORMÁTICA, 14., 2013, Campos do Jordão. **Anais [...]** São José dos Campos: INPE, 2013. p. 19-30.

CHEATHAM, M.; KRISNADHI, A.; AMINI, R.; HITZLER, P.; JANOWICZ, K.; SHEPHERD, A.; NAROCK, T.; JONES, M.; JI, P. The GeoLink knowledge graph. **Big Earth Data**, London, v. 2, n. 2, p. 131–143, 2018. DOI: 10.1080/20964471.2018.1469291.

DE MENEZES, L. C.; BAPTISTA, C. de S.; FALCÃO, A. G. R.; DE OLIVEIRA, M. G.; MARINHO, L. B. DYSCS: A platform to build geographically and semantically enhanced social content sites. **Journal of Systems and Software**, New York, v. 94, p. 39–49, 2014. DOI: 10.1016/j.jss.2014.03.059.

DO, B.-L.; TRINH, T.-D.; WETZ, P.; KIESLING, E.; ANJOMSHOAA, A.; TJOA, A. M. Multiscale exploration of spatial statistical datasets: a linked data mashup approach. *In*: INTERNATIONAL WORKSHOP ON SEMANTIC STATISTICS, 2., 2014, Riva del Garda. **Proceedings [...]** Riva del Garda: CEUR, 2014. Não paginado. Disponível em: <http://ceur-ws.org/Vol-1550/article-08.pdf>. Acesso em: 24 fev. 2021.

FERNÁNDEZ-ORDÓÑEZ, Y. M.; MEDINA-RAMÍREZ, R. C.; SORIA-RUIZ, J. Geographic metadata and ontology-based satellite image management. *In*: 2014 IEEE GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 2014, Quebec City. **Proceedings [...]** New York: IEEE, 2014. p. 117-120. DOI: 10.1109/IGARSS.2014.6946370.

FUGAZZA, C.; PEPE, M.; OGGIONI, A.; TAGLIOLATO, P.; CARRARA, P. Raising semantics-awareness in geospatial metadata management. **ISPRS International Journal of Geo-Information**, Basel, v. 7, n. 9, p. 370, 2018. DOI: 10.3390/ijgi7090370.

FUGAZZA, C.; PEPE, M.; OGGIONI, A.; TAGLIOLATO, P.; CARRARA, P. Streamlining geospatial metadata in the Semantic Web. **IOP Conference Series: Earth and Environmental Science**, Bristol, v. 34, p. 012009, 2016. DOI: 10.1088/1755-1315/34/1/012009.

FUGAZZA, C.; PEPE, M.; OGGIONI, A.; TAGLIOLATO, P.; PAVESI, F.; CARRARA, P. Describing Geospatial Assets in the Web of Data: A Metadata Management Scenario. **ISPRS International Journal of Geo-Information**, Basel, v. 5, n. 12, p. 229, 2016. DOI: 10.3390/ijgi5120229.

FUGAZZA, C.; TAGLIOLATO ACQUAVIVA D'ARAGONA, P.; OGGIONI, A.; CARRARA, P. Decentralized geospatial metadata management. **Earth Science Informatics**, Dublin, [s. v., s. n., s. p.], 2021. DOI: 10.1007/s12145-020-00559-1. Disponível em: <https://link.springer.com/article/10.1007/s12145-020-00559-1>. Acesso em: 24 fev. 2021.

FUGAZZA, C.; TAGLIOLATO, P.; FRIGERIO, L.; CARRARA, P. Web-Scale Normalization of Geospatial Metadata Based on Semantics-Aware Data Sources. **ISPRS International Journal of Geo-Information**, Basel, v. 6, n. 11, p. 354, 2017. DOI: 10.3390/ijgi6110354.

GIANNOPOULOS, G.; MAROULIS, T.; SKOUTAS, D.; KARAGIANNAKIS, N.; ATHANASIOU, S. FAGI-tr: a tool for aligning geospatial RDF vocabularies. *In*: ESWC 2014 SATELLITE EVENTS, 2014, Cham. **Proceedings [...]** Cham: Springer International Publishing Switzerland, 2014. p. 355-361. DOI: 10.1007/978-3-319-11955-7_48.

GRAUPNER, A.; NÜST, D. Serverless GEO labels for the semantic sensor web. *In*: INTERNATIONAL CONFERENCE ON GEOGRAPHIC INFORMATION SCIENCE, 11., 2020, Dagstuhl. **Proceedings [...]** Dagstuhl: Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020. p. 4:1-4:14. DOI: 10.4230/LIPIcs.GIScience.2021.I.4.

JIE YUAN; PENG YUE; JIANYA GONG; MINGDA ZHANG. A Linked Data Approach for Geospatial Data Provenance. **IEEE Transactions on Geoscience and Remote Sensing**, New York, v. 51, n. 11, p. 5105–5112, 2013. DOI: 10.1109/TGRS.2013.2249523.

KALANTARI, M.; RAJABIFARD, A.; OLFAT, H.; WILLIAMSON, I. Geospatial Metadata 2.0 – An approach for Volunteered Geographic Information. **Computers, Environment and Urban Systems**, New York, v. 48, p. 35–48, 2014. DOI: 10.1016/j.compenvurbsys.2014.06.005.

LI, Z.; CHIANG, Y.-Y.; TAVAKKOL, S.; SHBITA, B.; UHL, J. H.; LEYK, S.; KNOBLOCK, C. A. An automatic approach for generating rich, linked geo-metadata from historical map images. *In*: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY & DATA MINING, 26., 2020, New York. **Proceedings** [...] New York: ACM, 2020. p. 3290-3298. DOI: 10.1145/3394486.3403381.

MCKEE, G. The Map as a Search Box: Using Linked Data to Create a Geographic Discovery System. **Information Technology and Libraries**, Chicago, v. 38, n. 1, p. 40–52, 2019. DOI: 10.6017/ital.v38i1.10592.

PAGANI, R. N.; KOVALESKI, J. L.; RESENDE, L. M. Methodi Ordinatio: a proposed methodology to select and rank relevant scientific papers encompassing the impact factor, number of citations, and year of publication. **Scientometrics**, Amsterdam, v. 105, n. 3, p. 2109–2135, 2015. DOI: 10.1007/s11192-015-1744-x.

PEREIRA, C. M.; FERNEDA, E.; SANTAREM SEGUNDO, J. E. Análise do processo de recuperação da informação em bases de dados publicadas como dados abertos ligados utilizando a abordagem RDB2LOD. **Em Questão**, Porto Alegre, v. 26, n. 3, p. 94–120, 2020.

PO, L.; BIKAKIS, N.; DESIMONI, F.; PAPASTEFANATOS, G. Linked Data Visualization: Techniques, Tools, and Big Data. **Synthesis Lectures on the Semantic Web: Theory and Technology**, San Rafael, v. 10, n. 1, p. 1–157, 2020. DOI: 10.2200/S00967ED1V01Y201911WBE019.

RENTERIA-AGUALIMPIA, W.; LÓPEZ-PELLICER, F. J.; LACASTA, J.; ZARAZAGA-SORIA, F. J.; MURO-MEDRANO, P. R. Identifying hidden geospatial resources in catalogues. *In*: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON WEB INTELLIGENCE, MINING AND SEMANTICS, 3., 2013, New York. **Proceedings** [...] New York: ACM, 2013. p. 1-7. DOI: 10.1145/2479787.2479812.

RITSCHER, B.; MENDE, V.; GERICKE, L.; KORNMESSE, R.; PFEIFFER, S. Web Approach for Ontology-Based Classification, Integration, and Interdisciplinary Usage of Geoscience Metadata. **Data Science Journal**, London, v. 11, p. S140–S153, 2012. DOI: 10.2481/dsj.IGY-014.

SHVAIKO, P.; FARAZI, F.; MALTESE, V.; IVANYUKOVICH, A.; RIZZI, V.; FERRARI, D.; UCELLI, G. Trentino government linked open geo-data: a case study. *In*: INTERNATIONAL SEMANTIC WEB CONFERENCE, 11., 2012, Boston. **Proceedings** [...] Berlin: Springer, 2012. p. 196-211. DOI: 10.1007/978-3-642-35173-0_13.

W3C. WORLD WIDE WEB CONSORTIUM; OGC. OPEN GEOSPATIAL CONSORTIUM. Spatial Data on the Web Best Practices. W3C Working Group Note. 2017. Disponível em: <https://www.w3.org/TR/sdw-bp/#bp-metadata>. Acesso em: 10 maio 2021.

YUE, P.; GONG, J.; DI, L.; HE, L. Automatic geospatial metadata generation for earth science virtual data products. **Geoinformatica**, Boston, v. 16, n. 1, p. 1–29, 2012.

ZHANG, M.; YUAN, J.; GONG, J.; YUE, P. An interlinking approach for linked geospatial data. *In*: INTERNATIONAL ARCHIVES OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES, 2013, Antalya. **Proceedings** [...] Hannover: Copernicus GmbH, 2013. p.283-287. DOI: 10.5194/isprsarchives-XL-7-W2-283-2013.